

CLAIMS

What is claimed is:

1. A method implemented on a gateway device to provide load balancing gateway services to hosts on a network segment, the method comprising:
 - receiving an ARP message from a host addressed to an address shared by a plurality of gateway devices available for serving the hosts on the network segment;
 - based on load balancing considerations, selecting one of the plurality of gateway devices to act as the addressee gateway device for the host; and
 - replying to the ARP message with a reply message identifying the selected addressee gateway device.
2. The method of claim 1 wherein the gateway devices are layer 3 devices.
3. The method of claim 1 wherein the address shared by the plurality of gateway devices is a virtual IP address.
4. The method of claim 1 wherein the reply message identifies a layer 2 address for the addressee gateway device.
5. The method of claim 4 wherein the layer 2 address for the addressee gateway device is a virtual (secondary) MAC address.
6. The method of claim 1 further comprising the steps of:
 - determining that one of the plurality of gateway devices has failed; and
 - compensating for the failed gateway device by taking over acting as the addressee gateway device for any hosts for which the failed gateway device had been serving as the addressee gateway device.

7. The method of claim 1 further comprising the step of determining that the failed gateway device has failed permanently and redistributing addressee gateway device responsibility to other available gateway devices in the plurality of gateway devices for any hosts for which the failed gateway device had been serving as the addressee gateway device.

8. A method, implemented on a single gateway device, providing gateway services to hosts on a network segment, the method comprising:

receiving an ARP message from a host;

based on load balancing considerations, selecting a gateway device, from among a group of available gateway devices available for servicing hosts on the network segment, to act as the addressee gateway device for the host that sent the ARP message;

replying to the ARP message with a reply message identifying the addressee gateway device;

determining that one of the available gateway devices has failed; and

taking over responsibility for gateway services of the failed gateway device.

9. The method of claim 8 wherein the gateway device on which the method is implemented is a layer 3 device.

10. A computer-readable medium containing programmed instructions arranged to be implemented on a gateway device to provide load balancing gateway services to hosts on a network segment, the computer-readable medium including programmed instructions for:

receiving an ARP message from a host addressed to an address shared by a plurality of gateway devices available for serving the hosts on the network segment;

based on load balancing considerations, selecting one of the plurality of gateway devices to act as the addressee gateway device for the host; and

replying to the ARP message with a reply message identifying the selected addressee gateway device.

11. The computer-readable medium of claim 10 wherein the gateway devices are layer 3 devices.

12. The computer-readable medium of claim 10 wherein the address shared by the plurality of gateway devices is a virtual IP address.

13. The computer-readable medium of claim 10 wherein the reply message identifies a layer 2 address for the addressee gateway device.

14. The computer-readable medium of claim 13 wherein the layer 2 address for the addressee gateway device is a virtual (secondary) MAC address.

15. The computer-readable medium of claim 10 further comprising programmed instructions for:
determining that one of the plurality of gateway devices has failed; and
compensating for the failed gateway device by taking over acting as the addressee gateway device for any hosts for which the failed gateway device had been serving as the addressee gateway device.

16. The computer-readable medium of claim 10 further comprising programmed instructions for determining that the failed gateway device has failed permanently and redistributing addressee gateway device responsibility to other available gateway devices in the plurality of gateway devices for any hosts for which the failed gateway device had been serving as the addressee gateway device.

17. A computer-readable medium containing programmed instructions arranged to be implemented on a gateway device to provide load balancing gateway services to hosts on a network segment, the computer-readable medium including programmed instructions for:

receiving an ARP message from a host;

based on load balancing considerations, selecting a gateway device, from among a group of available gateway devices available for servicing hosts on the network segment, to act as the addressee gateway device for the host that sent the ARP message;

replying to the ARP message with a reply message identifying the addressee gateway device;

determining that one of the available gateway devices has failed; and

taking over responsibility for gateway services of the failed gateway device.

18. The computer-readable medium of claim 17 wherein the gateway device on which the method is implemented is a layer 3 device.

19. A master gateway device for use in a gateway load balancing service comprising the master gateway device and a slave gateway device, the master gateway device comprising:

a memory configured to hold:

a shared address shared by the master gateway device and the slave gateway device;

a plurality of unique addresses comprising:

a unique address for the master gateway device; and

a unique address for the slave gateway device;

a network interface configured to send and receive network traffic including an ARP message from a host addressed to the shared address;

a processor configured to:

select, based on load balancing considerations, one of the master and slave gateway devices to act as the addressee gateway device for the host; and

reply to the ARP message with a reply message identifying the selected addressee gateway device.

20. The device of claim 19 wherein the master and slave gateway devices are layer 3

devices.

21. The device of claim 19 wherein the address shared by the master and slave gateway devices is a virtual IP address.
22. The device of claim 19 wherein the reply message identifies a layer 2 address for the addressee gateway device.
23. The device of claim 22 wherein the layer 2 address for the addressee gateway device is a virtual (secondary) MAC address.
24. The device of claim 19 wherein the processor is further configured to:
detect the failure of one of the gateway devices; and
reconfigure the master gateway device to take over acting as the addressee gateway device for any hosts for which the failed gateway device had been serving as the addressee gateway device.
25. The device of claim 19 wherein the processor is further configured to:
detect whether the failed gateway device has failed permanently; and
if the failed gateway device has failed completely, to redistribute addressee gateway device responsibility to other available gateway devices in the plurality of gateway devices for any hosts for which the failed gateway device had been serving as the addressee gateway device.
26. A gateway apparatus to provide load balancing gateway services to hosts on a network segment, the apparatus comprising:
means for receiving an ARP message from a host addressed to an address shared by a plurality of gateway devices available for serving the hosts on the network segment;
means for selecting, based on load balancing considerations, one of the plurality of gateway devices to act as the addressee gateway device for the host; and
means for replying to the ARP message with a reply message identifying the

selected addressee gateway device.

27. The apparatus of claim 26 wherein the gateway devices are layer 3 devices.

28. The apparatus of claim 26 wherein the address shared by the plurality of gateway devices is a virtual IP address.

29. The apparatus of claim 26 wherein the reply message identifies a layer 2 address for the addressee gateway device.

30. The apparatus of claim 29 wherein the layer 2 address for the addressee gateway device is a virtual (secondary) MAC address.

31. The apparatus of claim 26 further comprising:
means for determining that one of the plurality of gateway devices has failed; and
means for compensating for the failed gateway device by taking over acting as the addressee gateway device for any hosts for which the failed gateway device had been serving as the addressee gateway device.

31. The apparatus of claim 26 further comprising means determining that the failed gateway device has failed permanently and means for redistributing addressee gateway device responsibility to other available gateway devices in the plurality of gateway devices for any hosts for which the failed gateway device had been serving as the addressee gateway device.

32. A gateway apparatus providing gateway services to hosts on a network segment, the apparatus comprising:

means for receiving an ARP message from a host;

means for selecting, based on load balancing considerations, a gateway device, from among a group of available gateway devices available for servicing hosts on the

network segment, to act as the addressee gateway device for the host that sent the ARP message;

means for replying to the ARP message with a reply message identifying the addressee gateway device;

means for determining that one of the available gateway devices has failed; and

means for taking over responsibility for gateway services of the failed gateway device.

33. The apparatus of claim 32 wherein the gateway apparatus is a layer 3 device.

34. A network segment comprising:

a plurality of gateway devices sharing an address, including a first gateway device;

a host configured to send an ARP message to the shared address one of the plurality of gateway devices;

wherein one of the gateway devices is configured to respond to the ARP message by sending the host a reply message identifying an addressee gateway device in the plurality of gateway devices.

35. The network segment of claim 34 wherein the plurality of gateway devices is configured to assume responsibility for any addressee gateway device that fails.

36. A network segment comprising:

-- a host having means for addressing messages to a layer 3 addressee device;

-- a layer 3 device group including a plurality of layer 3 devices; and

-- means for distributing packets from the host to the devices in the device group, the distributing means comprising:

-- means for electing one of the devices a master device, the electing means comprising:

-- means for determining a priority status of each device in the device group, wherein the master device is elected on the basis of its priority status and

includes:

- means for assigning a device in the device group as the addressee device for packets from the host;
- fail-over means comprising:
 - slave failure means for accommodating failure of one of the devices that is not the master device in the device group, wherein the slave failure means comprises:
 - means for re-distributing packets destined for a failed device to a functioning device in the device group;
 - master failure means for accommodating failure of the master device, wherein the master failure means comprises:
 - means for appointing a successor master device from among the functioning devices in the device group;
 - means for assigning the successor master device as the addressee device for packets from the host destined for the failed master..

37. A network device comprising:
a network interface; and
a processor designed or configured to perform the following operations:
receiving an ARP message from a host addressed to an address shared by a plurality of gateway devices available for serving the hosts on the network segment;
based on load balancing considerations, selecting one of the plurality of gateway devices to act as the addressee gateway device for the host; and
replying to the ARP message with a reply message identifying the selected addressee gateway device.

38. The network device of claim 37, wherein the network device is a layer 3 device.

39. The network device of claim 37 wherein the address shared by the plurality of gateway devices is a virtual IP address.

40. The network device of claim 37 wherein the reply message identifies a layer 2

address for the addressee gateway device.

41. The network device of claim 40 wherein the layer 2 address for the addressee gateway device is a virtual (secondary) MAC address.

42. The network device of claim 37, wherein the processor is further designed or configured to perform the following operations:

determining that one of the plurality of gateway devices has failed; and

compensating for the failed gateway device by taking over acting as the addressee gateway device for any hosts for which the failed gateway device had been serving as the addressee gateway device.

43. The network device of claim 37, wherein the processor is further designed or configured to perform the following operation:

determining that the failed gateway device has failed permanently and redistributing addressee gateway device responsibility to other available gateway devices in the plurality of gateway devices for any hosts for which the failed gateway device had been serving as the addressee gateway device.